



HAL
open science

A comparison between a direct and a multigrid sparse linear solvers for highly heterogeneous flux computations

Anthony Beaudoin, Jean-Raynald de Dreuzy, Jocelyne Erhel

► To cite this version:

Anthony Beaudoin, Jean-Raynald de Dreuzy, Jocelyne Erhel. A comparison between a direct and a multigrid sparse linear solvers for highly heterogeneous flux computations. CFM 2007 - 18ème Congrès Français de Mécanique, Aug 2007, Grenoble, France. hal-03362361

HAL Id: hal-03362361

<https://hal.science/hal-03362361>

Submitted on 1 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparaison entre deux solveurs linéaires : direct et multigrille

Anthony Beaudoin¹, Jean Raynald de Dreuzy² & Jocelyne Erhel³

¹ LMPG - Université du Havre - anthony.beaudoin@univ-lehavre.fr

² Géosciences - Université de Rennes 1 - aupepin@univ-rennes1.fr

³ IRISA - INRIA de Rennes - joceylne.erhel@irisa.fr

Abstract :

Subsurface hydraulic properties are mainly governed by the heterogeneity of the porous medium considered. Our work aims at characterizing the asymptotic dispersion coefficients for highly heterogeneous permeability fields triggered by advection and constant local dispersion - diffusion. We have developed a fully parallel software for simulating flow and transport. We have compared two well known sparse linear solvers, based respectively on a multi frontal Cholesky factorization and an iterative structured multigrid method. We study the effect of various parameters : the system size, the number of processors and the degree of heterogeneity of the permeability field.

Résumé :

L'hétérogénéité des formations géologiques induit une très forte variation des paramètres hydrauliques. Notre objectif est de caractériser les coefficients asymptotiques de la dispersion pour des champs de perméabilité hautement hétérogènes. Nous avons développé un logiciel complètement parallélisé pour simuler l'écoulement et le transport de solutés inertes soumis à de la convection et de la dispersion cinématique. Nous avons comparé deux solveurs de systèmes linéaires creux, basés respectivement sur une factorisation de Cholesky multi frontale et sur une méthode itérative multigrille structurée. Nous étudions les effets des différents paramètres, principalement la taille du système linéaire, le nombre de processeurs et le degré d'hétérogénéité du champ de perméabilité.

Key-words :

sparse linear solvers, heterogeneous flux computations, macro dispersion

1 Introduction

Les expériences menées dans une large gamme de milieux géologiques ont révélé des caractéristiques communes au transport de solutés tels que l'effet de l'hétérogénéité géologique, l'augmentation de la dispersion avec l'échelle et la lente convergence vers un régime asymptotique. La détermination de l'effet de l'hétérogénéité est un pré requis à la détermination des lois de transport et à l'établissement des règles de changement d'échelle (de l'échelle de l'observation géologiques à l'échelle de la modélisation hydrologique). Les expériences de terrain étant limitées en taille et en résolution, ce sont les simulations numériques de type Monte-Carlo qui permettent de comprendre et de quantifier ces règles de changement d'échelle. Numériquement, l'enjeu est de résoudre les problèmes de flux et de transport sur un grand nombre de milieux de très grandes tailles avec une très bonne résolution aboutissant à des systèmes linéaires de plusieurs dizaines de millions d'inconnues. Des progrès récents sur les solveurs parallèles et l'architecture des calculateurs parallèles permettent de résoudre ces difficultés. Nous utilisons ces progrès pour développer un logiciel de calcul parallèle. Le calcul des flux revient à résoudre un système linéaire caractérisé par une matrice creuse symétrique définie positive. La résolution de ce système linéaire est obtenue en utilisant deux bibliothèques. PSPASES et HYPRE sont deux bibliothèques offrant respectivement un solveur parallèle basé sur une factorisation de Cholesky multi frontale et un solveur parallèle basé sur une méthode itérative multigrille struc-

turée. Nous étudions les effets de différents paramètres sur ces deux solveurs, principalement la taille du système linéaire, le nombre de processeurs et le degré d'hétérogénéité du champ de perméabilité. Le papier est organisé de la manière suivante : nous décrivons d'abord le modèle et nous réalisons ensuite des tests de performance.

2 Modèle physique et numérique

Les simulations numériques Monte Carlo sont obtenues à partir de la répétition successive de quatre étapes : génération d'un champ aléatoire de perméabilité, évaluation du champ eulérien écoulement, identification de la trajectoire des particules et détermination des moments statistiques (Salandin *et al.* (1998)).

2.1 Champ de perméabilité

Le milieu poreux est supposé isotrope et caractérisé par une perméabilité aléatoire K (en m/s) qui suit une distribution logarithmique $Y = \ln(K)$, définie par une moyenne m_y (en m/s) et une fonction de covariance C_y :

$$C_y(\mathbf{r}) = \sigma_y^2 \exp(-|\mathbf{r}|^2/\lambda_y) \quad (1)$$

σ_y^2 et λ_y désignent la variance et la longueur de corrélation. La longueur de corrélation λ_y est comprise entre 0.1 m et 100 m (Tompson *et al.* (1998)); et la variance σ_y^2 entre 0 et 7 (Neuman *et al.* (1993)). \mathbf{r} est un vecteur de séparation (en m) entre deux vecteurs positions \mathbf{X} et \mathbf{X}' . Le champ de perméabilité est généré avec une simulation spectrale appelée FIM (Fourier Integral Method) (Yao (2004)). La simulation spectrale détermine la transformée de Fourier avec la librairie FFTW (Fast Fourier Transform in the West) (Frigo *et al.* (2005)) qui permet d'évaluer une transformée de Fourier discrète sur un cluster de processeurs.

2.2 Champ d'écoulement

Etant permanent et sans sources ni puits, l'écoulement est décrit en couplant l'équation de conservation de masse avec la loi de Darcy :

$$\nabla \cdot V = 0 \quad \text{avec} \quad V = -K\nabla h \quad (2)$$

V et h désignent la vitesse de Darcy (en m/s) et la charge hydraulique (en m). Les équations d'écoulement sont discrétisées sur une grille régulière en utilisant une méthode de volumes finis, permettant de conserver localement la masse et les flux sur les éléments. Pour les conditions aux limites, on impose une condition de Neumann $\partial h/\partial n = 0$ sur les frontières supérieure et inférieure. Sur les autres frontières, on impose deux conditions de Dirichlet $h = 0$ et 1 m. Le système linéaire $Ax = B$ est caractérisé par une matrice creuse A , définie positive et symétrique. La solution x est la charge hydraulique et le second membre B représente les conditions aux limites. Nous résolvons ce système avec la librairie HYPRE contenant un algorithme parallèle multigrille V-cycle (Falgout *et al.* (2005)) ou avec la librairie PSPASES contenant un algorithme parallèle de factorisation de Cholesky multi frontale (Gupta *et al.* (1999)).

2.3 Trajectoire des particules

Nous considérons le cas de la migration d'un soluté inerte dans un milieu poreux saturé. Cette migration est gouvernée par l'advection, la diffusion moléculaire et la dispersion cinématique.

Pour simplifier l'étude, les effets de la diffusion moléculaire et de la dispersion cinématique sont négligés dans un premier temps. Ce problème de transport est résolu en utilisant une méthode lagrangienne appelée méthode particulaire. Dans cette représentation lagrangienne, le problème de transport est résolu en déterminant la trajectoire de particules :

$$\frac{dX_i}{dt} = U_i \quad (3)$$

X_i est le vecteur position de la particule P_i . Cette méthode n'introduit pas de diffusion numérique pour une convection pure. Pour l'équation précédente, on utilise un schéma d'Euler explicite d'ordre 1. Le pas de temps dt (en secondes) est proportionnel au temps convectif, égale à la taille d'une maille divisée par le minimum des vitesses calculées sur une maille. A l'intérieur d'une maille, la vitesse des particules est obtenue par une interpolation linéaire respectant la conservation de la masse (Pollock (1988)). Ce modèle Transport est facilement parallélisable puisque les particules inertes n'interagissent pas entre elles. La difficulté réside dans l'accès à des données distribuées sur les processeurs. Gérant un sous domaine de la grille, chaque processeur calcule la trajectoire de la particule présente dans son sous domaine. Il communique aux processeurs voisins les particules sortantes de son sous domaine par des opérations, send et recv, non bloquantes et assurées par la librairie MPI.

2.4 Moments statistiques

Dans le cas d'un soluté inerte, la théorie du transport montre que les moments de la concentration sont déterminés à partir des moments caractérisant la distribution des particules (Dentz (2000)). Une forme discrète de ces moments est donnée par :

$$\langle x^k(t) \rangle_j = \frac{1}{N_p} \sum_{i=1}^{N_p} x_i(t)_j^k \quad (4)$$

N_p , j et i sont respectivement le nombre de particules injectées, le numéro de la simulation et le numéro de la particule injectée. k désigne l'ordre du moment calculée (1 ou 2). $x_i(t)$ est l'abscisse de la particule P_i . Les coefficients de dispersion sont calculés en prenant la dérivée temporelle de ces moments :

$$D_L(t) = \frac{1}{M} \sum_{j=1}^M \frac{d(\langle x^2(t) \rangle_j - \langle x(t) \rangle_j^2)}{dt} \quad (5)$$

M est le nombre de simulations réalisées. Pour obtenir la macro dispersion transversale, il suffit de remplacer L et x par T et y dans les expressions.

3 Tests numériques

Nous étudions l'effet de deux grandeurs physiques : la dimension du domaine L_x et la variance σ_y^2 . La dimension L_x varie de 256 à 4096. La variance σ_y^2 est comprise entre 0,5 et 6. Tous les tests sont réalisés sur un cluster SUN composé de deux noeuds de 32 processeurs chacun. Chaque processeur est un bi-processeur AMD Opteron 2,2 Ghz avec 2 Go de RAM. Ce cluster est un élément de la ressource informatique GRID'5000 de l'INRIA de Rennes.

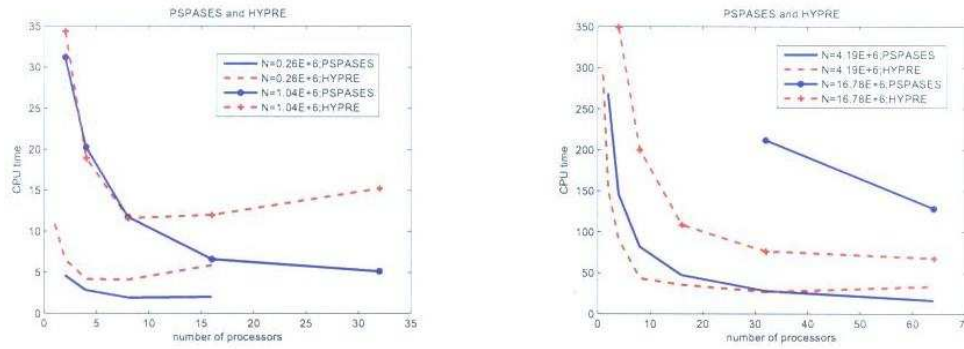


Figure 1: Temps CPU (en secondes) en fonction du nombre de processeurs P pour des systèmes linéaires de tailles différentes (en rouge HYPRE et en bleu PSPASES).

3.1 Analyse des performances

3.1.1 Temps CPU parallèle

Pour ces premiers tests de performance du logiciel, nous fixons la variance $\sigma_y^2 = 1$ et la dimension $512 \leq L_x \leq 2048$. Le nombre de processeurs P varie entre 2 et 32. Les temps CPU (en secondes) sont représentés sur la figure fig. 1. Pour un domaine de petite taille, PSPASES est plus efficace que HYPRE quelque soit le nombre de processeurs. Pour un domaine de taille moyenne, les deux solveurs donnent des temps similaires pour un nombre donné de processeurs. Pour un nombre important de processeurs, PSPASES devient plus rapide. Pour des domaines de grande taille, HYPRE est le plus rapide avec peu de processeurs mais devient lent lorsqu'on augmente le nombre de processeurs. Pour un domaine de grande taille, PSPASES est limité par l'espace mémoire disponible sur les machines.

3.1.2 Accélération

Nous avons également représenté sur la figure fig. 2 l'accélération en fonction du nombre de processeurs P . Les données sont les mêmes que celles du test précédent. Comme PSPASES doit être exécuté sur un nombre minimum de 2 processeurs, l'accélération est prise égale à $2T(2)/T(P)$ où $T(P)$ est le temps (en secondes) avec P processeurs. L'accélération augmente avec la taille du domaine parce que les deux solveurs semblent scalables. Cependant PSPASES a une meilleure efficacité que celle de HYPRE. Cela explique pourquoi les courbes des temps CPU se coupent (voir fig. 1). Pour PSPASES, nous avons également calculé l'efficacité en évaluant le ratio $R = (N1.5/(PT(P)))$, qui est approximativement proportionnel à l'efficacité $E = T(1)/PT(P)$. Nous observons que R reste constant lorsqu'on augmente la taille du domaine en maintenant constant le ratio N/P . Cela montre que PSPASES est bien scalable.

3.2 Influence de l'hétérogénéité

Nous étudions maintenant l'effet de très fortes hétérogénéités en variant la variance σ_y^2 entre 0.5 et 6. Nous gardons la taille du système linéaire N et le nombre de processeurs P fixes. Sur la figure fig. 3 gauche, nous avons représenté le nombre de V cycles générés par le solveur HYPRE. Il augmente très rapidement de 10 à 1000. 1000 est la valeur limite imposée. Sur la figure fig. 3 droite, nous observons également une augmentation rapide du temps CPU (en secondes) pour HYPRE. Avec PSPASES, le temps CPU reste constant. Cela montre que la vari-

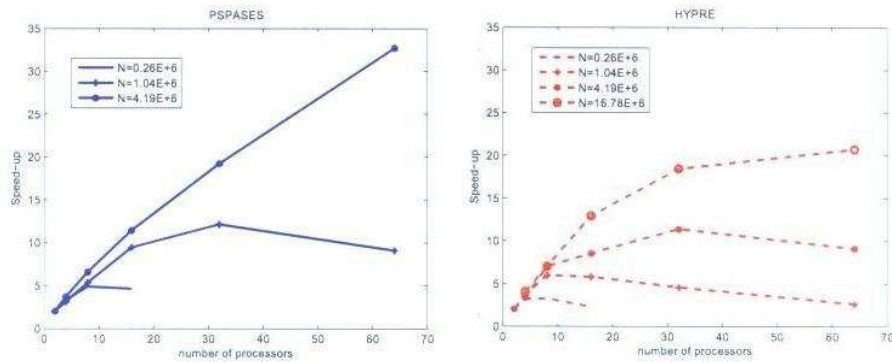


Figure 2: Accélération en fonction du nombre de processeurs P pour des systèmes linéaires de tailles différentes (en rouge HYPRE et en bleu PSPASES).

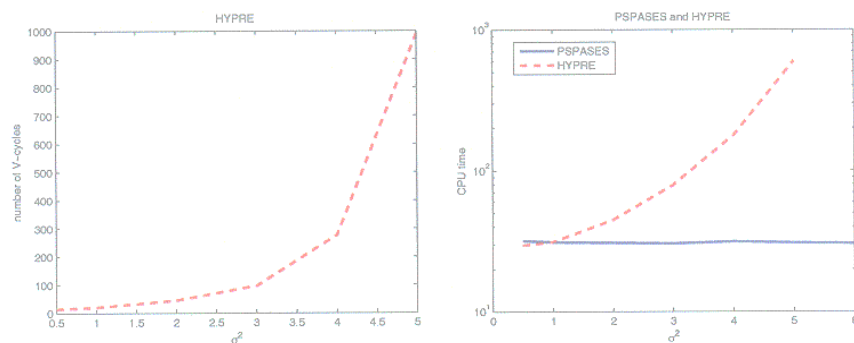


Figure 3: Nombre de V-cycles (figure de gauche) et temps CPU (figure de droite) en fonction de la variance σ_y^2 (en rouge : HYPRE et en bleu : PSPASES).

ance σ_y^2 n'a pas d'effets sur la performance de PSPASES. Pour PSPASES, nous avons observé des résidus très faibles. Mais comme le nombre de conditionnement augmente avec la variance, nous pouvons nous attendre à une perte de précision sur la solution. Pour une variance croissante, la méthode multigrille est donc moins efficace que la méthode directe. Cependant la méthode directe nécessite des temps CPU et un espace mémoire plus importants.

4 Conclusion

A l'aide de ce logiciel parallèle, nous pouvons simuler l'écoulement et le transport dans des milieux poreux 2D très fortement hétérogènes de très grande taille. Dans nos simulations, PSPASES apparaît être très efficace pour des domaines de tailles petites et moyennes alors que HYPRE est plus rapide pour des domaines de grandes tailles. HYPRE est très sensible au degré d'hétérogénéité qui fait augmenter le nombre de conditionnement. HYPRE devient alors moins efficace que PSPASES pour des milieux fortement hétérogènes. Une étude plus approfondie est à réaliser pour mieux comprendre le comportement de ces deux solveurs. Ce logiciel va nous être très utile pour déterminer les paramètres de dispersion dans des milieux poreux hétérogènes. Cette étude est en cours. Nous souhaitons également programmer une extension 3D de ce logiciel et lancer des simulations Monte Carlo 3D sur différents clusters connectés sur une grille de calcul.

References

- Dentz M., Kinzelbach H., Attinger S., Kinzelbach W. 2000 Temporal behaviour of a solute cloud in a heterogeneous porous medium 1 : point-like injection. *Water Resources Research* **36** 3591-3604
- Falgout, R.D., Jones, J.E., Yang U.M. 2005 Pursuing scalability for hypre's conceptual interfaces. *ACM Trans. Math. Softw.* **31** 326-350
- Frigo, M., Johnson, S.G. 2005 The design and implementation of FFTW3. *Proceedings of the IEEE* **93** 216-231
- Gupta, A., Gustavson, F., Joshi, M., Karypis, G., Kumar, V.. 1999 PSPASES : An efficient and scalable parallel sparse direct solver. In parallel numerical computations with applications. *T. Yang (ed.), Kluwer international series in engineering and computer science* **515**
- Neuman, S.P., Orr, S. 1993 Prediction of steady state flow in nonuniform geologic media by conditional moments : exact nonlocal formalism, effective conductivities and weak approximation. *Water Resources Research* **29** 341-364
- Pollock, D.W. 1998 Semianalytical computation of path lines for finite difference models. *Ground Water* **26** 743-750
- Salandin, P., Fiorotto, V. 1998 Solute transport in highly heterogeneous aquifers *Water Resources Research* **39**
- Tompson, A., Falgout, R.D., Smith, S.G., Bosl, W.J., Ashby S.F. 1998 Analysis of subsurface contaminant migration and remediation using high performance computing. *Advances in Water Resources* **22** 203-221
- Yao, T. 2004 Reproduction of the mean, variance and variogram model in spectral simulation. *Mathematical Geology* **36** 487-505